

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2001, 10:35:39 ; Search time 28.78 Seconds  
(without alignments)  
2378.577 Million cell updates/sec

Title: US-09-497-967-7  
Perfect score: 2540  
Sequence: 1 MKNILVILISLFINQIKS.....QCDPANELSLLSLLISYLL 468

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues 473505  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_17:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2540	100.0	468	5 Q9BMH3	Q9bmh3 ichthyophth
2	921	36.3	442	5 Q9XZG2	Q9xzg2 ichthyophth
3	775.5	30.5	394	5 Q27208	Q27208 ichthyophth
4	345	13.6	371	5 Q9GPP0	Q9gpp0 tetrahymena
5	289.5	11.4	316	5 Q9GPP3	Q9gpp3 tetrahymena
6	286.5	11.3	316	5 Q9GPP4	Q9gpp4 tetrahymena
7	283.5	11.2	305	5 Q9GPP2	Q9gpp2 tetrahymena
8	236	9.3	594	5 Q24970	Q24970 giardia lam
9	232	9.1	645	5 Q97448	Q97448 giardia lam
10	228	9.0	1275	5 Q24977	Q24977 giardia lam
11	227.5	9.0	667	5 Q9XTK3	Q9xtk3 giardia lam
12	226	8.9	719	5 Q9U019	Q9u019 giardia lam
13	221.5	8.7	719	5 Q9U021	Q9u021 giardia lam
14	219	8.6	1274	5 Q9NGL3	Q9ngl3 giardia lam
15	218.5	8.6	548	5 Q9GQ45	Q9gq45 caenorhabdi
16	218.5	8.6	1372	5 P91526	P91526 caenorhabdi
17	218	8.6	1154	5 Q9GQ46	Q9gq46 giardia lam
18	216.5	8.5	436	5 Q27197	Q27197 tetrahymena
19	216.5	8.5	597	5 Q07317	Q07317 giardia int

20	210.5	8.3	560	5	Q9U013	Q9u013 giardia lam
21	210.5	8.3	769	5	Q24971	Q24971 giardia lam
22	209	8.2	704	5	Q9U048	Q9u048 giardia lam
23	206	8.1	3396	5	Q9VM55	Q9vm55 drosophila
24	205	8.1	709	5	Q9XTJ7	Q9xtj7 giardia lam
25	202	8.0	1551	5	Q9NGV4	Q9ngv4 drosophila
26	200	7.9	1704	5	Q94446	Q94446 chironomus
27	199	7.8	709	5	Q97444	Q97444 giardia lam
28	198	7.8	394	5	Q9GQ47	Q9gq47 giardia lam
29	197	7.8	557	5	Q24992	Q24992 giardia lam
30	196	7.7	503	5	Q9U018	Q9u018 giardia lam
31	193	7.6	556	5	Q9NGZ3	Q9ngz3 giardia lam
32	191.5	7.5	704	3	Q74567	Q74567 trichoderma
33	191.5	7.5	1297	5	Q26632	Q26632 strongyloce
34	190.5	7.5	2972	5	P90891	P90891 caenorhabdi
35	189	7.4	1622	5	Q06550	Q06550 cryptospori
36	189	7.4	5374	11	Q99ND0	Q99nd0 mus musculu
37	188.5	7.4	350	5	Q94589	Q94589 lembadion b
38	188.5	7.4	2759	5	Q45614	Q45614 caenorhabdi
39	188.5	7.4	3102	5	Q9TZR4	Q9tzzr4 caenorhabdi
40	188	7.4	739	5	Q9GS24	Q9gs24 giardia lam
41	187.5	7.4	1459	5	O17084	O17084 caenorhabdi
42	187	7.4	423	5	Q9U697	Q9u697 tetrahymena
43	186	7.3	545	5	Q9GQ44	Q9gq44 giardia lam
44	184.5	7.3	3567	11	Q9ES77	Q9es77 mus musculu
45	183.5	7.2	1679	5	Q24301	Q24301 drosophila

ALIGNMENTS

RESULT 1  
Q9BMH3  
ID Q9BMH3 PRELIMINARY; PRT: 468 AA.  
AC Q9BMH3;  
DT 01-JUN-2001 (T-EMBLrel. 17, Created)  
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)  
DE IMMOBILIZATION ANTIGEN ISOFORM.  
GN IAG52A.  
OS Ichthyophthirius multifiliis.  
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;  
OC Ophryoglenina; Ichthyophthirius.  
OX NCBI\_TaxID=5932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=G5;  
RA Lin Y., Lin T.-L., Clark T.G.;  
RT "Variation in primary sequence and tandem repeat copy number among i-  
RT antigen genes of Ichthyophthirius multifiliis.";  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF324424; AAK01661.1; -;  
SQ SEQUENCE 468 AA; 48281 MW; BEA6DA42833A7726 CRC64;

Query Match 100.0%; Score 2540; DB 5; Length 468;  
Best Local Similarity 100.0%; Pred. No. 3e-192;  
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKNILVILISLFINQIKSANGCPVGTETNTAGQVDDLGPANCVCNKFNYYNNAAFV 60  
|||||  
Db 1 MKNILVILISLFINQIKSANGCPVGTETNTAGQVDDLGPANCVCNKFNYYNNAAFV 60  
|||||

QY 61 PGASTCTPCQKKDAGAQPNPPATANLVTCQNVKCPAGTAIAGGATDYAAIITECVNCRI 120  
|||||

Db 61 PGASTCTPCQKKDAGAQPNPPATANLVTCQNVKCPAGTAIAGGATDYAAIITECVNCRI 120  
|||||

QY 121 NFYENAPNFNAGASTCTACPNVRVGGALTAGNAATIVACQNVACPTGTALDDGVTTDYV 180  
|||||

Db 121 NFYENAPNFNAGASTCTACPNVRVGGALTAGNAATIVACQNVACPTGTALDDGVTTDYV 180  
|||||

QY 181 RSPTECVKRLNFYYNNGNNGTPTFNPCKSQCTPCPAIKPANVAQATLGNDAITTAQCNTVA 240  
|||||

118 DVFDRSAAQCCKKPNFYNGSGPOGEAPGVGVFAAGAAAAGVAAVTSCQVPCOLNK--N 172  
 QY 222 VAOATLGNDAITAQCNVACPDGTGTSIAAGVGNVNWVAQNTSE---CTNCAPNFYNN-----N 272  
 Db 176 DSPATAGAQAANLATQCSNQCPGTGVLDDGVT--LVFNTSATLCVKRPNFYNGSGPOGE 233  
 QY 273 APN---FNPQ-----NSTCLPCPANKDYGAEATAGGATATLAKOCNIACTPDGTAIAS 320  
 Db 234 APGVGVFAAGAAAAGVAAVTSCQVPCQINKN--DSPATAGAQAANLATQCSQCPGTGTAIQD 292  
 QY 321 GAT--NYVILQTECLNCAANFYFGNNFNOAGSRRCACPANKVQGVAVATAGGTATLIAQCA 379  
 Db 293 GVTLVFNSNSTQCSQCIANFFNG--NFEAGKSQCLKCPVSKTTPAHA--PGNTATQATQCL 350  
 QY 380 LECPAAGTIVLDTGTTSTYKQAASECVCVCAANFYTTKTDVWAGIDTCTSCNKKLTSAGEAN 439  
 Db 351 TTCPCAGTIVLDDGTSTNFVASATEKCSAGFASKTTGTAGTDTCTCTCKLTSGATAK 410  
 QY 440 LPESAKKNIQO---DFANFLSISLILSYVLL 468  
 Db 411 VYAEATQKVQCASTTFKAFUSISLILFSFYLL 442  
  
 RESULT 3  
 Q27208  
 ID Q27208 PRELIMINARY; PRT; 394 AA.  
 AC Q27208;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
 DE 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
 DE IMMobilization ANTIGEN PRECURSOR (FRAGMENT).  
 OS Ichthyophthirius multifiliis.  
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida  
 OC Eukaryogliena; Ichthyophthirius.  
 OX NCBI\_TaxID=5932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GEORGIA;  
 RX MEDLINE=92335298; PubMed=1631132;  
 RA Clark T.G., McGraw R.A., Dickerson H.W.;  
 RT "Developmental expression of surface antigen genes in the parasitic  
 RT ciliate Ichthyophthirius multifiliis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6363-6367(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GEORGIA;  
 RX MEDLINE=93020590; PubMed=1383510;  
 RA Lin T.L., Dickerson H.W.;  
 RT "Purification and partial characterization of immobilization antigen  
 RT from Ichthyophthirius multifiliis.";  
 RL J. Protozool. 39:457-463(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GEORGIA;  
 RA Clark T.;  
 RN Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GEORGIA;  
 RA Clark T.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M92907; AAC36158.1; -.  
 KW Signal.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 1 POTENTIAL.  
 SQ SEQUENCE 394 AA; 39495 MW; 3013C2B2BEFDB682 CRC64;  
  
 Query Match 30.5%; Score 775.5; DB 5; Length 394;  
 Best Local Similarity 39.8%; Pred. No. 1.9e-53;  
 Matches 182; Conservative 36; Mismatches 128; Indels 111; Gaps  
 23 CPVGTENTAGQVD----DLGTFPANCVCQRNFYNNAAAFVPGASTCTPCPKKDKAGQ 78

```
Db 3 CPDGTQF-AGLTDVCAADLGT---CVNCRPNFYNGGAA-----QGEAN 43
QY 79 PNPATANLVTCNVKCPAGTAAGATDYAAIITECVNCRINFYNAPNPNAGASTCT 138
Db 44 GNOPEAAN-----NAARGICV 59
QY 139 ACPVNRVGCALTAGNAATIAOQNVACPTGTALDGVYTDYVRSFTECVKRLNRYNGN 198
Db 60 PCQINRVGVTNAGDLATLQCTQCPTGTALDGVTDVDFRSAAQCCKPKNFYNGG 119
QY 199 N--GNTP-----RSQCTPCPAIKPANVAQATLGNDAITTAOCNVACPD 243
Db 120 SPQGEAPGVQVFAAGAAAAGVAATVSQCVPCLNK--NDSPATAGAQAQNLATQCSNQCP 177
QY 244 GTISAAGVNNWVAQNT---CTNCAPNFYNN-----NAPN---FNPG-----NST 282
Db 178 GTVLDGVNT--LVFNSTSLVCVKCRPNFYNGGSPQGEAPGVQVFAAGAAAAGVAATVSQ 235
QY 283 CLPCPANKDYGAEATAGGAATLAKOCNACPDGTAIASGAT--NYVILQTECLNCAANFYF 341
Db 236 CVPCLNKN--DSPATAGAQAQNLATQCTQCPTGTALDGVYTDVDFRSAAQCCKPKNFYNGG 294
QY 342 DGNFQAASSRCKACPAKVKQAVAGTAGTATLIAQCALECPAGTVLTDGTTSTYKQAAS 401
Db 295 NG--NLEAGKSQLCKPVSKTTPAHA--PGNTATQATQCLTTCPAGTVLTDGTTSTYKQAAS 352
QY 402 ECVKCAANFYTKQTDWAGIDTCTSCNKKLTSGAEA 438
Db 353 ECTKCSAGFFASKTGTGTAGTDTCTECTKLTSGAEA 389
```

```
RESULT 4
Q9GPP0 PRELIMINARY; PRT; 371 AA.
AC Q9GPP0;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE IMMOBILIZATION ANTIGEN LD (FRAGMENT).
GN SERLB.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ANF18211;
RX MEDLINE=20549003; PubMed=11095959;
RA Doerder F.P., Gerber C.A.;
RT "Molecular Characterization of the SerL Paralogs of Tetrahymena thermophila."
RL Biochem. Biophys. Res. Commun. 278:621-626(2000).
DR EMBL: AF312775; AAC38107.1; -.
FT NON_TER 1
SQ SEQUENCE 371 AA; 35175 MW; 5817EFFC2517DEAC CRC64;
```

```
Query Match 13.68; Score 345; DB 5; Length 371;
Best Local Similarity 30.18; Pred. No. 1.3e-19;
Matches 141; Conservative 41; Mismatches 175; Indels 112; Gaps 33;
```

```
QY 8 ILIISLFINQIKSAN--CPVGTETTAGQVDDLTGPANVCNCKNFYNNAAAFVPGASTC 66
Db 6 LILISLAV--IATVNAC---TDTNATA-----GAGGTCTF--CNAGYGTSTDTVTASGA--C 52
QY 67 TPCQKQKADAGQPNPPATANLVTCNVKCPAGTAAGATDYAAIITECVNCRINFYNEN 126
Db 53 QKCPGTGNSVA---ATASGLTVISCT-----CNDTNAGLKADNSG-----CQCKANFY--G 98
QY 127 APNENAGAST--CTACPNRVGGALTAGNAATIAOQNVACPTGTALDGVYTDYVRSFTE 185
Db 99 TPNVAVAGGTGCTACP---TGTASPGATAAATVSCACN-----DTNASLKGDN 143
```

```
QY 186 CVKRLNRYNGNNGTTPFNPCKSQCTPCPAIKPANVAQATLGNDAITTAOCNVACPDGT 245
Db 144 GCCKANFYGTGN---AVAGGATGCTACP-----TGSAAAAGSTAVTSCACN-----DT 189
QY 246 ISAAGVNNWVAQNTCTNCAPNFYNNAPNPNPGNST--CLPCPANKDYGAEATAGGAATL 304
Db 190 NSAL-----KADNSACI--CKANFY--GTPNAVAGGATGCTACPT---GSAAAAAGSTAVT 237
QY 305 AKOCNIAICPDGATATASGATNVVILQTECLNCAANFYFDGNNFQAGSSRCKACPAKVKQA 364
Db 238 SCACN-----DTNSALKADN-----SACI--CKANFYGTPNAVAGGATGCTACPT---GT 282
QY 365 VATAGGTATLIAQCALECPAGTVLTDGTTSTYKQAASECVKCAANFYTTTQTDWAGIDT 424
Db 283 TSTAG--TTVIGSCA--CP-----DTNASLNTATPPVQCCKANFYGTPTTTGASG--- 328
QY 425 CTSCNKKLTSGAEANLPESAKKNIOQCFAN-----FLSISLLISYLL 468
Db 329 CTAC-----PSQTA--PAGSATNV--CKAASSTSYILPIVLSLLFLSLVML 370
```

```
RESULT 5
Q9GPP3 PRELIMINARY; PRT; 316 AA.
AC Q9GPP3;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE IMMOBILIZATION ANTIGEN LB (FRAGMENT).
GN SERLB.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B;
RX MEDLINE=20549003; PubMed=11095959;
RA Doerder F.P., Gerber C.A.;
RT "Molecular Characterization of the SerL Paralogs of Tetrahymena thermophila."
RL Biochem. Biophys. Res. Commun. 278:621-626(2000).
DR EMBL: AF312771; AAG38117.1; -.
FT NON_TER 1
SQ SEQUENCE 316 AA; 30008 MW; 133A0B7D0797A3BD CRC64;
```

```
Query Match 11.4%; Score 289.5; DB 5; Length 316;
Best Local Similarity 30.1%; Pred. No. 2.6e-15;
Matches 118; Conservative 35; Mismatches 140; Indels 99; Gaps 29;
```

```
QY 8 ILIISLFINQIKSAN--CPVGTETTAGQVDDLTGPANVCNCKNFYNNAAAFVPGASTC 66
Db 6 LILISLAV--IATVNAC---TDTNATA-----GAGGTCTF--CNAG--YGTSTDTVTPSGS--C 52
QY 67 TPCQKQKADAGQPNPPATANLVTCNVKCPAGTAAGATDYAAIITECVNCRINFYNEN 126
Db 53 TKCPGTGNSAA---ATASGLTVSSCT--CNDTNASLKGDN-----SGC--QCKANFY--G 98
QY 127 APN--FNAGASTCTACPNRVGGALTAGNAATIAOQNVACPTGTALDGVYTDYVRSFTE 185
Db 99 TPNVAVAGGTGCTACP---TGTTSPTAGTAATVSCACN-----DTNASLKGDN 143
QY 186 CVKRLNRYNGNNGTTPFNPCKSQCTPCPAIKPANVAQATLGNDAITTAOCNVACPDGT 245
Db 144 GCCKANFYGTGN---AVSGGATGCTACPTGSAA-----AAGSTAVTSCACNDTN 190
QY 246 ISAAGVNNWVAQNTCTNCAPNFYNNAPNPNPGNST--CLPCPANKDYGAEATAGGAATL 304
Db 191 SSLK-----ADNSACV--CKANFY--GTPNAVAGGATGCTACPT---GTTSTAG--TTV 235
QY 305 AKOCNIAICPDGTAIASGATNVVILQTECLNCAANFYFDGNNFQAGSSRCKACPAKVKQA 364
```

```

Db 236 IGSC--ACPDNALNSATPPV-----CQKANFY--GPTASGASCTACPS-----GQ 281
QY 365 VATAGGTATLIAQCALECPAGTAVLTGTTSTY 396
Db 282 TAPA-GSATNVCKAA-----STSSY 301

RESULT 6
Q9GPP4 PRELIMINARY; PRT; 316 AA.
AC Q9GPP4;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE IMMOBILIZATION ANTIGEN LA (FRAGMENT).
GN SERLA.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B;
RX MEDLINE=20549003; PubMed=11095959;
RA Doerder F.P., Gerber C.A.;
RT "Molecular Characterization of the SerL Paralogs of Tetrahymena
thermophila.";
RL Biochem. Biophys. Res. Commun. 278:621-626(2000).
DR EMBL; AF312770; AAG38116.1; -.
FT NON_TER 1
FT SEQUENCE 316 AA; 30100 MW; 1A13D076F28ED3BD CRC64;
SQ

Query Match 11.3%; Score 286.5; DB 5; Length 316;
Best Local Similarity 30.1%; Pred. No. 4.5e-15;
Matches 118; Conservative 35; Mismatches 140; Indels 99; Gaps 29;

QY 8 ILIISLFINOIKSAN-CPVGTETNTAGOVDDLGTTPANCVCQKNFYNNAAAFVPGASTC 66
Db 6 LILISLAV--IATVNAC---TDTNATA-----GAGTCTCF-CNAG-YGTSTDTVTPSGS-C 52
QY 67 TPCPQKKDAGAPNPATANLVTCNVKCPAGTAAGATDYAAIITEVCNCRINFYNEN 126
Db 53 TKCPTGTNSAA--ATASGTLVSSCT--CNDTNASLKGDN-----SGC-QCKANFY--G 98
QY 127 APN-FNAGASTCTACPNRVGGALTAGNAATIVAQCNAVACPTGTALDDGVTTDYVRSFTE 185
Db 99 TPNNAVSGGATGCSACP---TGTTSPAGTAAVTSCACN-----DTNASLKGDN 143
QY 186 CVKCRLFNYNGNNTFPNPGKSOCTPCPAIKPANVAQATLGNDAITIAQCNAVACPDGT 245
Db 144 GCQCKANFYGTNP-----AVSGGATGCTACTGSA-----AAGSTAVTSCACNDTN 190
QY 246 ISAAGVNNWAQNTCTNCAPNFYNNNAPNPNST-CLPCPANKDYGAETAGGAATL 304
Db 191 SSLK-----ADNSACV-CKANFY--GTPNAVAGGATGCTACTP-----GTTSTAG--TTV 235
QY 305 AKQCNACPDGTATASGATNVVILQTECLNCAANFYDGNFQAGSSRCKACPAKPVQGA 364
Db 236 IGSC--ACPDNALNSATPPV-----CQKANFY--GPTASGASCTACPS-----GQ 281
QY 365 VATAGGTATLIAQCALECPAGTAVLTGTTSTY 396
Db 282 TAPA-GSATNVCKAA-----STSSY 301

RESULT 7
Q9GPP2 PRELIMINARY; PRT; 305 AA.
AC Q9GPP2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)

```

```

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE IMMOBILIZATION ANTIGEN LC (FRAGMENT).
GN SERLC.
OS Tetrahymena thermophila.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymenina; Tetrahymena.
OX NCBI_TaxID=5911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B;
RX MEDLINE=20549003; PubMed=11095959;
RA Doerder F.P., Gerber C.A.;
RT "Molecular Characterization of the SerL Paralogs of Tetrahymena
thermophila.";
RL Biochem. Biophys. Res. Commun. 278:621-626(2000).
DR EMBL; AF312772; AAG38118.1; -.
FT NON_TER 1
FT SEQUENCE 305 AA; 28863 MW; 0568C353A0253564 CRC64;
SQ

Query Match 11.2%; Score 283.5; DB 5; Length 305;
Best Local Similarity 30.1%; Pred. No. 7.6e-15;
Matches 118; Conservative 34; Mismatches 141; Indels 99; Gaps 29;

QY 8 ILIISLFINOIKSAN-CPVGTETNTAGOVDDLGTTPANCVCQKNFYNNAAAFVPGASTC 66
Db 6 LILISLAV--IATVNAC---TDTNATA-----GAGTCTCF-CNAG-YGTSTDTVTPSGS-C 52
QY 67 TPCPQKKDAGAPNPATANLVTCNVKCPAGTAAGATDYAAIITEVCNCRINFYNEN 126
Db 53 TKCPTGTNSVA--ATASGTLVSSCT--CNDTNASLKGDN-----SGC-QCKANFY--G 98
QY 127 APN-FNAGASTCTACPNRVGGALTAGNAATIVAQCNAVACPTGTALDDGVTTDYVRSFTE 185
Db 99 TPNNAVSGGATGCSACP---TGTTSPAGTAAVTSCACN-----DTNASLKGDN 143
QY 186 CVKCRLFNYNGNNTFPNPGKSOCTPCPAIKPANVAQATLGNDAITIAQCNAVACPDGT 245
Db 144 GCQCKANFYGTNP-----AVAGGATGCTACTGSA-----AAGSTAVTSCACNDTN 190
QY 246 ISAAGVNNWAQNTCTNCAPNFYNNNAPNPNST-CLPCPANKDYGAETAGGAATL 304
Db 191 SSLK-----ADNSACV-CKANFY--GTPNAVAGGATGCTACTP-----GTTSTAG--TTV 235
QY 305 AKQCNACPDGTATASGATNVVILQTECLNCAANFYDGNFQAGSSRCKACPAKPVQGA 364
Db 236 IGSC--ACPDNALNSATPPV-----CQKANFY--GPTASGASCTACPS-----GQ 281
QY 365 VATAGGTATLIAQCALECPAGTAVLTGTTSTY 396
Db 282 TAPA-GSATNVCKAA-----STSSY 301

RESULT 8
Q24970 PRELIMINARY; PRT; 594 AA.
AC Q24970;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE VARIANT-SPECIFIC SURFACE PROTEIN.
GN VSPA6-SL.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WBAG;
RA Yang Y., Adam R.D.;
RL Infect. Immun. 0:0-0(0).
DR EMBL; U17980; AAA82585.1; -.
InterPro; IPR000345; CytC_heme_bind.

```

```
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002174; Furin-like.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR SMART: SM00181; EGF; 2.
DR SMART: SM00261; FU; 6.
SQ SEQUENCE 594 AA; 59575 MW; 9680818FB75F52AC CRC64;

Query Match          9.3%; Score 236; DB 5; Length 594;
Best Local Similarity 23.2%; Pred. No. 7.9e-11;
Matches 131; Conservative 47; Mismatches 207; Indels 180; Gaps 32;

QY 20 SANCPCVGTETAGQVDDLTGTPANCVCQKNF--YNNAAAFVPGASTCTPCPKKADAGAQ 78
Db 40 NGNTPLYLKKTNPS--DPTGTCVSAVDCGSGAGYITDSD--VSDAKECKKC----- 85

QY 79 PNPP-----ATANLVTC-----NVKCPAGTAIA-----GGA-----TDYAAII 112
Db 86 -NAPCTACAGTADKCTCKDANGAAPYLKKNPSDPTGCVSAVDCGSGAGYITDSDVSDA 144

QY 113 TECVNCRINFYENAPN-----ENAGASTCTACPNRVGGALTAGNAATIVACQNVACP 166
Db 145 KECKKA----EQKPNTAGTCQFCSDANCERCQNDVCARCSTG-APPENGRCPAATP 199

QY 167 TGTALDDGVTTDYVRSFTE-CYKCRNLNFIN-----GNNG-----NTPP----- 204
Db 200 GCHSSCDGCTENAMTNOADKCTGCKEGRYLKPESAAGSGTCLTAECTSDTHTEKEA 259

QY 205 NFGKSOCTPCPAIKPA-----NVAQATLGNATITTAQCNVACPQGTISAAQNNVNAQNT 259
Db 260 GDSKGMLPCSDATHGACGCKKALKTLGSAESTVVGCS-ECTDKWLTPSG-----NA 311

QY 260 ECTNCAPNFINNAPNFNPGNSTCLPCPANKD-----YCAEATAGGAATLA 305
Db 312 CLDNCPCAGTYPNDNLCTSCHDTCAECGNADRASCTACYPGYSLLYS-CTAG---TCV 367

QY 306 KQNTA-----CPDG-----TAI-ASGATNY-VILQTECL 333
Db 368 KECTGAFGANCADGCTADVGGAKYCAQCKDGYAPIDGICTAVAAAGRTNVCTRAADGCT 427

QY 334 NCAANF-YFDGNF-----QAQSRCKACPAKPVGGAVATAGGATLIAQAL 380
Db 428 KCAGETLTMGGCGYVAKLPKSVCTLASNCKCINCAAN-----GQAPVQEKCP- 476

QY 381 ECPAG-----TVLTDGTTSTYKQAASECVKCAANFYTTTKQTDWAGIDTC----- 425
Db 477 ECSEGCACKNDNACTECLPGYIKGAGDKCFKCTASSGNNNQITGVANCVTCAPPAGSG 536

QY 426 -TSCNKKL--TSGAEANLPESAKN 447
Db 537 SVTCYVKTGDTSGDDNTGGSVNKS 561

RESULT 9
O97448 PRELIMINARY; PRT; 645 AA.
AC O97448;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CYSTEINE-RICH PROTEIN.
GN CRP65.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96186899; PubMed=8635746;
RA Chen N., Upcroft J.A., Upcroft P.;
RT "A new cysteine-rich protein-encoding gene family in Giardia
RL duodenalis."
RL Gene 169:33-38(1996).
DR EMBL; L39804; AAB06228.1; -.
```

```
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002174; Furin-like.
DR PROSITE: PS01186; EGF_2; UNKNOWN_4.
DR SMART: SM00261; FU; 3.
SQ SEQUENCE 645 AA; 65262 MW; F19FE98DB0AA589 CRC64;

Query Match          9.1%; Score 232; DB 5; Length 645;
Best Local Similarity 22.9%; Pred. No. 1.8e-10;
Matches 117; Conservative 40; Mismatches 183; Indels 170; Gaps 28;

QY 19 KSANPCVGTETAGQVDDLTGTPANCVCQKNFYNNAAAFVPGASTCTPC--PKKDAQ 76
Db 176 KASNTDGEELKRA-----CATC-----TAVGPNQOTCLTCNGGQKQVOLN 216

QY 77 AOPNPATANLVTCNVKCPAGTAIAGATDYAAIITECVNCRINFYN-----ENAPNFN 131
Db 217 G-----ISGDSGSPNSAANPD-----ICEC-NEGFNLNSGDKGCEKASNTD 257

QY 132 -----AGASTCTACPNRVGGALTAGNAATIV-----AOCNVACPTGTALDDGVTTDYV 180
Db 258 CGEELKRAAGCATCTAVGPN--GOTCLTCNGGQKQVOLNGISGDSGSPNSAANPD----- 310

QY 181 RSFTECVKCRNLNFINYNGN-----NTPFPNGKSQCTPCPAIKPAN--VAQAT 226
Db 311 -----CECNEGFNLNSGDKGCEKASNTQCTNP-----NCKICDNPKTDMEVCTECN 356

QY 227 LGNDATITTAQCNVACPQGTISAAQNNVNAQNTECTNCAPNFYNNNAPNFNPGNSTCLPC 286
Db 357 DGDYLTPTNQCVDPCT--TISGYGDN---DKKCKACSPC-----AECVGPANNQCSSC 406

QY 287 PANK--DYGAETAGGAATLAKQCNACPDGTATAS-----GATNYVI-----LOTE----- 331
Db 407 PAGKLLTYTDDSNPNNGGTCGDACKVSA-DGTGCTCGAQIGGTAYCCKTSTQAPLNG 465

QY 332 -----CLNCAANFYF-DG-----NNFOAGSSRCKACPA 358
Db 466 DCAASSRATCTKMGNGVCVQCEDNFFLKDGCGYQYDTRQPKQVCSNAQNGNGKQCTC--- 533

QY 359 NKVOGAVATAGGATLIAQALCEPCAGTTLTDGTTSTYKQAASECVKCAANFYTTTKQTDW 418
Db 524 --ANGLAATDGNCAECHPTCA-----TCSAPSTASSCKTCATGYIKENGDDT 568

QY 419 VAGIDTCTCNKKLTSCAEANLPESAKNI 448
Db 569 TDG--PCMCKSEKI-SGCKQCVSSGSSVI 595

RESULT 10
O24977 PRELIMINARY; PRT; 1275 AA.
AC O24977;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CYSTEINE RICH PROTEIN.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen N., Upcroft P., Upcroft J.;
RL Parasitology 111:0-0(0).
DR EMBL; L29079; AAA74587.1; -.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR000561; EGF-like.
DR PROSITE: PS002174; Furin-like.
DR SMART: SM00181; EGF; 23.
SQ SEQUENCE 1275 AA; 135848 MW; 58E21622872A971A CRC64;
```

Query Match 9.0%; Score 228; DB 5; Length 1275;

Best Local Similarity 23.1%; Pred. No. 7.1e-10;

Matches 119; Conservative 35; Mismatches 183; Indels 178; Gaps 30;

```
QY 21 ANCPV-GTETNTAGQVDDLTGTPANCVCNCKNEYYN-----NAAAFVPGASTCTPC 69
DB 801 APCNVEGCTCEVGNAAQ-----QCKTCRPGYTINTDTKQCTKDPKPEAPCNVEGCTCV-- 852
QY 70 POKKDAQAQP-----NPPATANLVTO-----CNVK-CPAGTAIAGATDYAAIITEC 115
DB 853 -----EGNAQCKTCRPGYTINTDTKQCTKDPKPEAPCNVEGCE--TCVEGNA-----QOC 899
QY 116 VNCRIIFYENAPNFNAGASTCTACPVNRVGCALTAGNAATIVACNVACPTGTALDDGV 175
DB 900 KTCRPGY-----TINTDTKQCTKDP-----EAPCNV-----EGC 928
QY 176 TTDYVRSFTECVKCRNLNFYNGNG-----NTPENPKGSKQCTPCPAIKPANVAQA 225
DB 929 ETCVEGNAQCKTCRPGYTINTDTKQCTKDPKPEAPCNTP-----NCKTCDNPKTDNEI-C 981
QY 226 TLGNDAITTAQCNVACPDGTISAAGV-----NNVQAQNTCTNCAPNFYNNAPNFNGN 280
DB 982 TKCNDGDLTPTNQCPDCT-AISGYIGYDTKKCKACNPECAEV-----GPN 1029
QY 281 STCLPCPANK--DYGAETAGGAATLAKQCNIA-----CPD-----GTAIASGATN--- 324
DB 1030 NQCTACPVGKMLQYTDNTFVNGGTCTMDQCSVSSSTNDGCAEGGAIGGTAICSKCKNTQQ 1089
QY 325 -----YVLOTELCNCAANFYF-DG-----NNFQAGSSRC 353
DB 1090 APLNGNCAASSRVAFCATITSGACTKCEGYFLKDGCGYQTDROPQKQVCSNAOAGNGKC 1149
QY 354 KACPANKVQAVATAGTATLIAQCALECPAGVTLNDGTTSTVKAASBCVKAANFYTT 413
DB 1150 QTC-----ANGLASDGNCA-ECHS-----TATCTST-ADRAADCKTCATGYIKE 1192
QY 414 KQTDWVAGIDTCTSCNKKLTSGAEANLPESAKNI 448
DB 1193 NGDDTTAGL--CKKCSKI-SGCKQCVSSSGSVI 1224
```

RESULT 11

```
Q9XTK3 ID Q9XTK3 PRELIMINARY; PRT; 667 AA.
AC Q9XTK3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VSP417-3/A-II.
DE VSP417-3/A-II.
GN Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIS-136;
RA Ey P.L., Darby J.M., Mayrhofer G.;
RX MEDLINE=99053029; PubMed=9836309;
RT "Comparison of tsal417-like variant-specific surface protein (VSP) genes in Giardia intestinalis and identification of a novel locus in genetic group II isolates.";
RL Parasitology 117:0-0(0).
RL EMBL; AF03384; AAD03497.1; -.
DR HSSP; P02468; ITLE.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM00261; FU; 3.
DR SMART; SM00261; FU; 3.
SQ SEQUENCE 667 AA; 69123 MW; 77C64CF59441C0C CRC64;
```

Query Match

9.0%; Score 227.5; DB 5; Length 667;

Best Local Similarity 24.3%; Pred. No. 4.1e-10;

Matches 120; Conservative 39; Mismatches 176; Indels 159; Gaps 33;

```
QY 18 IKSANCPVGTETNTAGQVDDLTGTPANCVN--CQKNFYNNAAAFV-----PG 62
DB 171 IKATATC-TGCDSNKIVKTDTSGT--SCIESASNGFFVSDQRKAQSKDYKICPRIDDDPA 227
QY 63 ASTCTPCPQKKDAGAPNPPATANLVTOCN-----VKCPAGTAIAGGAT 106
DB 228 KANTACSDNK-----KPNLEG-----TECNSTQDQHCFAEAGTCQKCSGFLDQ--- 275
QY 107 DYAAIITEC--VNCRIIFYENAPNFNAGASTCTAC-----PVNR-VGGALTAGN--A 154
DB 276 -QNCVKSCKTENCKA-----CTNPKAANEVCTECVFTHLTPTSOCVOYCTGLGNYA 328
QY 155 AT-----IVAQCQNV-CPTGTALDDGVTTDYVRSFTECVKCRNLNFYNGNGNTFPNG 207
DB 329 GTNADNKNACKECRVANCK--TCVDQG-----QCQTCNNGFYKNGDACSPCHESC 376
QY 208 K-----SQCTPCPAIKPANVAQATLGNDAITTAQCNVACPDGTISAAGVNNVVAQNT 260
DB 377 KTCAGTASDCTCKPTKALR-----YGNDBG--KGTCGEGCTTGKSGA----- 419
QY 261 CTNCAPNFYNNAPNFNGNSTCLPCPANKDYG-----AEATAGGAATLAKQCNIA--- 311
DB 420 CKTCGLTI-----DGASYCSECATTEYPONGICTSTTARTAT-CKNSNVANGIC 469
QY 312 --CPDGTAIASG---ATNY-----VILQT-----ECLNCAANFYFDG--NFOAGSSRCAC 356
DB 470 SLCTNGFLRMNGGCVETTKFPKSKVCTTTPDADCTSVTSYIIDGSSNLVVCSDGCAEC 529
QY 357 PANKVOGAVATAGGTATLI--AQALECPAGTAVLTGTTSTVKAASBCVKAANFYTTK 414
DB 530 TTS-----SACTCKDGTYTKIGNSQCTKQ-----DSCSETCTGAATCKACATGYTK- 578
QY 415 QTDWVAGIDTCTSC 428
DB 579 -----ALGESTCTTSC 588
```

RESULT 12

```
Q9U019 ID Q9U019 PRELIMINARY; PRT; 719 AA.
AC Q9U019;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VARIANT-SPECIFIC SURFACE PROTEIN.
GN VSP417-7.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRIS-136;
RA Ey P.L., Darby J.M.;
RT "A new locus (vsp417-7) belonging to the subfamily of tsal417-like variant-specific surface protein (vsp) genes in Giardia intestinalis.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067148; AAF21772.1; -.
DR HSSP; P35555; 1EMN.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00261; FU; 4.
DR SMART; SM00261; FU; 4.
SQ SEQUENCE 719 AA; 73874 MW; E409450249E3F716 CRC64;
```

Query Match

8.9%; Score 226; DB 5; Length 719;

Best Local Similarity 23.3%; Pred. No. 5.9e-10;

Matches 109; Conservative 40; Mismatches 178; Indels 140; Gaps 25;

```
QY 44 CVMCKNFY-----NNAAFVPGASTCTPCP-----QKDKAGAOBNPPATANLVTOCNV 93
Db 264 CDDTQKGFYKVDSTNGKNCVSCADGAGLAGVADGAKVGDCAKCTKPADINTPTKCD- 322
QY 94 KCPAGTAIAGGADYAAIITECV-----NCRINFYENAPFNAGASTCTACPVN----- 143
Db 323 ECKPGVEI---STDK-----TKCTSTAPDCPI-----ENCKVCSEDKRACEECNSNLYLTP 371
QY 144 -----RVGALTAGNA-----ATIVAQCNAVACPTGT---ALDDGVTTDYVRSFTE 185
Db 372 TRMCIDCKKIGNYTTNANKKLICKEAVANCEKCENTGTCKTCDG-----FYKSSSE 427
QY 186 CVKCRNFYNGNNGTNPENPKSOCTPCPAIKPANAQATLGNDAITIAQCNAVACPDGT 245
Db 428 CKACDSN---CKTCNGGT-----SADCTKC-----LSGAVLKVGNGT---KGTGAGCATGT 474
QY 246 ISAAGVNNVQAQNTNCTNCAFNFNAPNPNSTCLPCPANKDY-----GA 294
Db 475 GAGA-----CKTCG-----LIIDGTSYCECAVETEPQGGVCSSTTVRA 514
QY 295 EATAGGAATLAKOCNTACPDG-----TATASGATNVVILQTECL 333
Db 515 AATCKAGSVAKGMCN-SCTNGLFMRNGGCTTKFPKGSVCBEAASAGDT-----CQ 565
QY 334 NCAANFYFDGNNFOAGSSRCKACAPANKV-----OGAVATAGGTATLIAQCALECPAGTVL 388
Db 566 KEAPGYHLNNDLVTCSPGCKTCTSTVCTACMEGVKTSKCAKCAAGCA-TCTGSGTITA 624
QY 389 TCGTTSTYKQAASECVKCAANFYTTKQTD-WVAGIDTCTSCNKKILTS 434
Db 625 CDTCSGTGYESGTCVSC-----TESNSDKTITGVANCAASCAPPLNN 666

RESULT 13
Q9U021 ID Q9U021 PRELIMINARY; PRT; 719 AA.
AC Q9U021;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE VARIANT-SPECIFIC SURFACE PROTEIN.
GN VSP417-7.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AD-1;
RA Ey P.L., Darby J.M.;
RT "Comparative analysis of the VSP417 subfamily of variant-specific
proteins in Giardia intestinalis.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF189719; IEMN.
DR HSSP; P35555; 1EMN.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00001; EGF-like; 1.
DR SMART; SM00261; FU; 4.
DR SMART; SM00261; FU; 14.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 719 AA; 73888 MW; 83BE706BACE7F977 CRC64;
```

```
Query Match 8.7%; Score 221.5; DB 5; Length 719;
Best Local Similarity 23.1%; Pred. No. 1.3e-09;
Matches 113; Conservative 40; Mismatches 176; Indels 161; Gaps 28;

QY 27 TETNTAGQVDDL-----GTPAN-CVNCOKNFYNNAAAFV-----PGASTCTPCP 70
Db 256 TECVAANECDTQKGFYKVDSTNGKNCVSCA-----DGAGLAVGTDGAKVGDCAKCI 310
QY 71 QKDKAGAOBNPPATANLVTOCNVKAAGTAIAGGADYAAIITECV-----NCRINFYNE 125
Db 311 K-----PADINTPTKCD-ECKPGVEI---STDK-----TKCTSTAPDCPI-----E 348
```

```
QY 126 NAPFNAGASTCTACPVN-----RVGALTAGNA-----ATIVAQCNA 164
Db 349 NCKVCSEDKRACEECNSNLYLTPTRMCIDCKKIGNYTTNANKKLICKEAVANCEK 408
QY 165 CPTGT--ALDDGVTTDYVRSFTECVKCRNFYNGNNGTNPENPKSOCTPCPAIKPANV 222
Db 409 ENTGTCRCDG-----FYKSSSECKACDSN---CKTCNGGT-----SADCTKC-----LSG 452
QY 223 AOATLGNDAITIAQCNAVACPDGTISAAGVNNVQAQNTNCTNCAFNFNAPNPNST 282
Db 453 AVLKYGNDGT-KGTGAGCATGTGAGA-----CKTCG-----LIIDGTSY 491
QY 283 CLPCPANKDY-----GAETAGGAATLAKOCNTACPDG----- 315
Db 492 CSECAVETEPQGGVCSSTTVRAAATCKAGSVAKGMCN-SCTNGLFMRNGGCTTKFP 550
QY 316 -----TATASGATNVVILQTECLNCAANFYFDGNNFOAGSSRCKACAPANKV-----OGAV 365
Db 551 KSVCEEAASAGDT-----CQKEAPGYHLNNDLVTCSPGCKTCTSTVCTACMEGYV 602
QY 366 ATAGTATLIAQCALECPAGTVLTDGTTSTYKQAASECVKCAANFYTTKQTD-WVAGIDT 424
Db 603 KTSKCAKCAAGCA-TCTGTTTACDTCSTGYKSGTTCVSC-----TESNSDKTITGVAN 656
QY 425 CTSCNKKILTS 434
Db 657 CAXCAPPLNN 666

RESULT 14
Q9NGL3 ID Q9NGL3 PRELIMINARY; PRT; 1274 AA.
AC Q9NGL3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE VARIANT-SPECIFIC SURFACE PROTEIN VSP136B.
GN VSP136B.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AD-1;
RA Mansouri M., Ey P.L.;
RT "Analysis of a vsp136 homolog in Giardia intestinalis.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF249878; AAF69839.1; -.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002174; Furin-like.
DR SMART; SM00181; EGF; 23.
DR SMART; SM00261; FU; 14.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 1274 AA; 135671 MW; 68644A814BD6AE65 CRC64;
```

```
Query Match 8.6%; Score 219; DB 5; Length 1274;
Best Local Similarity 22.9%; Pred. No. 3.6e-09;
Matches 118; Conservative 35; Mismatches 184; Indels 178; Gaps 30;
```

```
QY 21 ANCPV-GTETNTAGQVDDLGTTPANCVCNOKNFYNN-----NAAAFVPGASTCTPC 69
Db 801 APCNVEGCTCCEGNAQ-----QCKTCRPGYTTINTDKQCTKDPKPEAPCNVEGCTCV-- 852
QY 70 PQKDKAGAOP-----NPPATANLVTO-----CNVK-CPAGTAIAGGADYAAIITEC 115
Db 853 -----EGNAQCKTCRPGYTTINTDKQCTKDPKPEAPCNVEGCE--TCVEGNA-----QQC 899
QY 116 VNCRINFYENAPFNAGASTCTACPVNVRVGGALTAGNAATIVAQCNAVCTGTALDDGV 175
Db 900 KTCRPGY-----TINTDTKQCTKDP-----EAPCNV-----EGC 928
```

QY 176 TTDYVRSFTECVKRLNFYNGNG-----NTPFNGKSOCTPCPAIKPANYAQA 225  
Db 929 ETCVGNQAQCKTCRPGYTINTDKQCTKDPFPCNTP-----NCKTCDNPKTDNEI-C 981  
QY 226 TLGNDATITACQNVACPDGTISAAGV-----NNVQAQNTCTCAPNFYNNAPNPN 280  
Db 982 TKCNDGDLTPTNQCPDCT-AISGYGDTDKKACNPECAECV-----GPN 1029  
QY 281 STCLPCPANK--DYGAETAGGAATLAKQCNIA-----CPD-----GTATAGATN--- 324  
Db 1030 NOCTACPVGKMLQYTDTPVNGGTCMDQCSVSTNDGCAEGAGIAGTAYCSKCKNTQ 1089  
QY 325 -----YVILQTECLNCAANFYF-DG-----NNFOAGSSRC 353  
Db 1090 APLNGNCAASSRVAFCATITSGACTKCEGYFLKDGCGYQTDROPQKQVCSNAQGGNGK 1149  
QY 354 KACPANKVOGAVATAGGTATLIAQALCEPAGTVLTDTTSTYKQAASECVKCAANFYTT 413  
Db 1150 QTC-----ANGLAASDGNCA-ECHS-----TCATCST-ADAADKCKTCTCATGYNKE 1192  
QY 414 KOTDWWAGIDTCTSCNKKLTSGAEANLPESAKNI 448  
Db 1193 NGDDTTAGL--CKKSEKI--SGCKQCVSSSGSVI 1224

RESULT 15  
Q9GQ45 PRELIMINARY; PRT; 548 AA.  
AC Q9GQ45;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE VARIANT-SPECIFIC SURFACE PROTEIN M21-1 (FRAGMENT).  
GN M21-1.  
OS Giardia lamblia (Giardia intestinalis).  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.  
OX NCBI\_TaxID=5741;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AD-1;  
RA Mansouri M., Ey P.L.;  
RT "A segment of a vsp72-like gene homolog from a type A-I (group 1)  
RT Giardia intestinalis isolate.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF298862; AAG37862.1; -  
DR InterPro; IPR000345; CyfC\_heme\_bind.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR002174; Furin-like.  
DR SMART; SM00181; EGF; 2.  
DR SMART; SM00001; EGF\_like; 1.  
DR SMART; SM00261; FU; 5.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
FT NON\_TER 1  
FT NON\_TER 548 548  
SQ SEQUENCE 548 AA; 56557 MW; 578FE4FDA0A2CF0E CRC64;

Query Match 8.6%; Score 218.5; DB 5; Length 548;  
Best Local Similarity 21.0%; Pred. No. 1.7e-09;  
Matches 123; Conservative 51; Mismatches 201; Indels 211; Gaps 32;  
QY 10 IISLFINQIKSANCPCVGTETNTAGQVDDLGTTPANCVCNCKNEFYNNAAAFVPG----- 62  
Db 3 LIPLFVVSALAVTC-----QADKCTVGNTEICTQCRARGVPVDFGFCWPPGFQAAAA 55  
QY 63 -----ASTCTPC-----PQKDAQAQNPANLVTQCNV----- 93  
Db 56 GCTEDGVPLDKTAATCGKGGDGLLFMGCGYKTESQPSGSDICTAASNGVCTECNTKGL 115  
QY 94 -KCPAGTATAG-----GATDYAAIITECVNCRINFYNENAPFNAGASTCTACP- 141  
Db 116 FRNPATAPEKGEICLCHDATGADGYMG-VEGCATC-----TAPTNNKGAATCTECOD 167

QY 142 -VNRVGA-----LTAGNAATIVA-----QCNVA--CPTGTALDDGVTTDFV 180  
Db 168 GYNDGGACKKCVDGICDCTGANOCCTCEDGKYLKNNOCVDAGCCDQGYADP--TTGQC 225  
QY 181 R--SFTECVKRLNFYNGNGNTPENPGKSOCTPCPAIKPANVAQATLGNDAITFAOCN 238  
Db 226 KPGGITDDCATCEYNATIS-----QPQCKTCTSSNKNMVKTA--DGTTCVDD 271  
QY 239 VACPDPG--TISAAGVN-----NWVAQNTCTCAPNFYNNAPNPNFG 279  
Db 272 GGCTNGNTHFVEGTNOKLCVPCGDTTNGGVLGNTCSTCTCTKCLDGYDGS-----G 326  
QY 280 NSTCLPCPANKDYGAETAGGAATLA-----KOCNIACPDGTATASGATNVVILQTECLNC 335  
Db 327 TVTCTACP-----GANCATLCERYKRQCT-TCKPGFFLKDSGG-----ECISC 369  
QY 336 AANFYDGNFQAGSSRCACAPKPVQGAATAG-----GTA-----TLIAQCALE--- 381  
Db 370 S-----DKNN--GGHEGCSACSSN--GAFKCTDCKPNYKKEGTSNNTCVKTCDEDETAC 419  
QY 382 -----CPAGTVLTDTSTYKQAASE-----CYKC-AANF 410  
Db 420 GGTSGACDAIVIDENGNTKHYCSFCGSGRFPIDGLCASDKANNNGCANGVCTCTAANY 479  
QY 411 YT-----TKQTDWVAGIDTCTSCNKKLTSGAEAN-----LPESAKIN 447  
Db 480 FLYMGCGYKNTVPGSHMCKTANNVCTAVSENKYPFIVPGASNQN 525

Search completed: December 26, 2001, 10:38:54  
Job time: 195 sec